

RECEIVED

MAY 16 2001

TECH CENTER 1600/2900

#1  
Dm  
5-11  
1652

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/619,047C

DATE: 05/11/2001  
TIME: 13:55:55

Input Set : A:\CHEM1110.ST25.txt  
Output Set: N:\CRF3\05112001\I619047C.raw

P.S

ENTERED

```

3 <110> APPLICANT: Chemicon International, Inc.
4     LENG, Jay
6 <120> TITLE OF INVENTION: PROTEASE SPECIFIC CLEAVABLE LUCIFERASES AND METHODS OF USE THEREOF
8 <130> FILE REFERENCE: CHEM1110
10 <140> CURRENT APPLICATION NUMBER: US 09/619,047C
11 <141> CURRENT FILING DATE: 2000-07-18
13 <160> NUMBER OF SEQ ID NOS: 29
15 <170> SOFTWARE: PatentIn version 3.0
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 936
19 <212> TYPE: DNA
20 <213> ORGANISM: Renilla reniformis
22 <220> FEATURE:
23 <221> NAME/KEY: CDS
24 <222> LOCATION: (1)..(936)
26 <400> SEQUENCE: 1
27 atg act tcg aaa gtt tat gat cca gaa caa agg aaa cgg atg ata act      48
28 Met Thr Ser Lys Val Tyr Asp Pro Glu Gln Arg Lys Arg Met Ile Thr
29 1           5           10           15
31 ggt ccg cag tgg tgg gcc aga tgt aaa caa atg aat gtt ctt gat tca      96
32 Gly Pro Gln Trp Trp Ala Arg Cys Lys Gln Met Asn Val Leu Asp Ser
33           20           25           30
35 ttt att aat tat tat gat tca gaa aaa cat gca gaa aat gct gtt att      144
36 Phe Ile Asn Tyr Tyr Asp Ser Glu Lys His Ala Glu Asn Ala Val Ile
37           35           40           45
39 ttt tta cat ggt aac gcg gcc tct tct tat tta tgg cga cat gtt gtg      192
40 Phe Leu His Gly Asn Ala Ala Ser Ser Tyr Leu Trp Arg His Val Val
41           50           55           60
43 cca cat att gag cca gta gcg cgg tgt att ata cca gat ctt att ggt      240
44 Pro His Ile Glu Pro Val Ala Arg Cys Ile Ile Pro Asp Leu Ile Gly
45 65           70           75           80
47 atg ggc aaa tca ggc aaa tct ggt aat ggt tct tat agg tta ctt gat      288
48 Met Gly Lys Ser Gly Lys Ser Gly Asn Gly Ser Tyr Arg Leu Leu Asp
49           85           90           95
51 cat tac aaa tat ctt act gca tgg ttt gaa ctt ctt aat tta cca aag      336
52 His Tyr Lys Tyr Leu Thr Ala Trp Phe Glu Leu Leu Asn Leu Pro Lys
53           100          105          110
55 aag atc att ttt gtc ggc cat gat tgg ggt gct tgt ttg gca ttt cat      384
56 Lys Ile Ile Phe Val Gly His Asp Trp Gly Ala Cys Leu Ala Phe His
57           115          120          125
59 tat agc tat gag cat caa gat aag atc aaa gca ata gtt cac gct gaa      432
60 Tyr Ser Tyr Glu His Gln Asp Lys Lys Ile Lys Ala Ile Val His Ala Glu
61           130          135          140
63 agt gta gta gat gtg att gaa tca tgg gat gaa tgg cct gat att gaa      480
64 Ser Val Val Asp Val Ile Glu Ser Trp Asp Glu Trp Pro Asp Ile Glu
65 145          150          155          160
67 gaa gat att gcg ttg atc aaa tct gaa gaa gga gaa aaa atg gtt ttg      528

```

## RAW SEQUENCE LISTING

DATE: 05/11/2001

PATENT APPLICATION: US/09/619,047C

TIME: 13:55:55

Input Set : A:\CHEM1110.ST25.txt

Output Set: N:\CRF3\05112001\I619047C.raw

```

68 Glu Asp Ile Ala Leu Ile Lys Ser Glu Glu Gly Glu Lys Met Val Leu
69      165      170      175
71 gag aat aac ttc ttc gtg gaa acc atg ttg cca tca aaa atc atg aga      576
72 Glu Asn Asn Phe Phe Val Glu Thr Met Leu Pro Ser Lys Ile Met Arg
73      180      185      190
75 aag tta gaa cca gaa gaa ttt gca gca tat ctt gaa cca ttc aaa gag      624
76 Lys Leu Glu Pro Glu Glu Phe Ala Ala Tyr Leu Glu Pro Phe Lys Glu
77      195      200      205
79 aaa ggt gaa gtt cgt cgt cca aca tta tca tgg cct cgt gaa atc ccg      672
80 Lys Gly Glu Val Arg Arg Pro Thr Leu Ser Trp Pro Arg Glu Ile Pro
81      210      215      220
83 tta gta aaa ggt ggt aaa cct gac gtt gta caa att gtt agg aat tat      720
84 Leu Val Lys Gly Gly Lys Pro Asp Val Val Gln Ile Val Arg Asn Tyr
85 225      230      235      240
87 aat gct tat cta cgt gca agt gat gat tta cca aaa atg ttt att gaa      768
88 Asn Ala Tyr Leu Arg Ala Ser Asp Asp Leu Pro Lys Met Phe Ile Glu
89      245      250      255
91 tcg gat cca gga ttc ttt tcc aat gct att gtt gaa ggc gcc aag aag      816
92 Ser Asp Pro Gly Phe Phe Ser Asn Ala Ile Val Glu Gly Ala Lys Lys
93      260      265      270
95 ttt cct aat act gaa ttt gtc aaa gta aaa ggt ctt cat ttt tcg caa      864
96 Phe Pro Asn Thr Glu Phe Val Lys Val Lys Gly Leu His Phe Ser Gln
97      275      280      285
99 gaa gat gca cct gat gaa atg gga aaa tat atc aaa tcg ttc gtt gag      912
100 Glu Asp Ala Pro Asp Glu Met Gly Lys Tyr Ile Lys Ser Phe Val Glu
101      290      295      300
103 cga gtt ctc aaa aat gaa caa taa      936
104 Arg Val Leu Lys Asn Glu Gln
105 305      310
108 <210> SEQ ID NO: 2
109 <211> LENGTH: 311
110 <212> TYPE: PRT
111 <213> ORGANISM: Renilla reniformis
113 <400> SEQUENCE: 2
115 Met Thr Ser Lys Val Tyr Asp Pro Glu Gln Arg Lys Arg Met Ile Thr
116 1      5      10      15
119 Gly Pro Gln Trp Trp Ala Arg Cys Lys Gln Met Asn Val Leu Asp Ser
120      20      25      30
123 Phe Ile Asn Tyr Tyr Asp Ser Glu Lys His Ala Glu Asn Ala Val Ile
124      35      40      45
127 Phe Leu His Gly Asn Ala Ala Ser Ser Tyr Leu Trp Arg His Val Val
128      50      55      60
131 Pro His Ile Glu Pro Val Ala Arg Cys Ile Ile Pro Asp Leu Ile Gly
132 65      70      75      80
135 Met Gly Lys Ser Gly Lys Ser Gly Asn Gly Ser Tyr Arg Leu Leu Asp
136      85      90      95
139 His Tyr Lys Tyr Leu Thr Ala Trp Phe Glu Leu Leu Asn Leu Pro Lys
140      100      105      110
143 Lys Ile Ile Phe Val Gly His Asp Trp Gly Ala Cys Leu Ala Phe His

```

## RAW SEQUENCE LISTING

DATE: 05/11/2001

PATENT APPLICATION: US/09/619,047C

TIME: 13:55:55

Input Set : A:\CHEM1110.ST25.txt

Output Set: N:\CRF3\05112001\I619047C.raw

```

144          115          120          125
147 Tyr Ser Tyr Glu His Gln Asp Lys Ile Lys Ala Ile Val His Ala Glu
148          130          135          140
151 Ser Val Val Asp Val Ile Glu Ser Trp Asp Glu Trp Pro Asp Ile Glu
152 145          150          155          160
155 Glu Asp Ile Ala Leu Ile Lys Ser Glu Glu Gly Glu Lys Met Val Leu
156          165          170          175
159 Glu Asn Asn Phe Phe Val Glu Thr Met Leu Pro Ser Lys Ile Met Arg
160          180          185          190
163 Lys Leu Glu Pro Glu Glu Phe Ala Tyr Leu Glu Pro Phe Lys Glu
164          195          200          205
167 Lys Gly Glu Val Arg Arg Pro Thr Leu Ser Trp Pro Arg Glu Ile Pro
168          210          215          220
171 Leu Val Lys Gly Gly Lys Pro Asp Val Val Gln Ile Val Arg Asn Tyr
172 225          230          235          240
175 Asn Ala Tyr Leu Arg Ala Ser Asp Asp Leu Pro Lys Met Phe Ile Glu
176          245          250          255
179 Ser Asp Pro Gly Phe Phe Ser Asn Ala Ile Val Glu Gly Ala Lys Lys
180          260          265          270
183 Phe Pro Asn Thr Glu Phe Val Lys Val Lys Gly Leu His Phe Ser Gln
184          275          280          285
187 Glu Asp Ala Pro Asp Glu Met Gly Lys Tyr Ile Lys Ser Phe Val Glu
188          290          295          300
191 Arg Val Leu Lys Asn Glu Gln
192 305          310
195 <210> SEQ ID NO: 3
196 <211> LENGTH: 936
197 <212> TYPE: DNA
198 <213> ORGANISM: Renilla reniformis (mutated sequence)
200 <220> FEATURE:
201 <221> NAME/KEY: CDS
202 <222> LOCATION: (1)..(936)
204 <400> SEQUENCE: 3
205 atg act tcg aaa gtt tat gat cca gaa caa agg aaa cgg atg ata act      48
206 Met Thr Ser Lys Val Tyr Asp Pro Glu Gln Arg Lys Arg Met Ile Thr
207 1          5          10          15
209 ggt ccg cag tgg tgg gcc aga tgt aaa caa atg aat gtt ctt gat tca      96
210 Gly Pro Gln Trp Trp Ala Arg Cys Lys Gln Met Asn Val Leu Asp Ser
211          20          25          30
213 ttt att aat tat tat gat tca gaa aaa cat gca gaa aat gct gtt att      144
214 Phe Ile Asn Tyr Tyr Asp Ser Glu Lys His Ala Glu Asn Ala Val Ile
215          35          40          45
217 ttt tta cat ggt aac gcg gcc tct tct tat tta tgg cga cat gtt gtg      192
218 Phe Leu His Gly Asn Ala Ala Ser Ser Tyr Leu Trp Arg His Val Val
219          50          55          60
221 cca cat att gag cca gta gcg cgg tgt att ata cca gat ctt att ggt      240
222 Pro His Ile Glu Pro Val Ala Arg Cys Ile Ile Pro Asp Leu Ile Gly
223 65          70          75          80
225 atg ggc aaa tca ggc aaa tct ggt aat ggt tct tat agg tta ctt gat      288

```

## RAW SEQUENCE LISTING

DATE: 05/11/2001

PATENT APPLICATION: US/09/619,047C

TIME: 13:55:55

Input Set : A:\CHEM1110.ST25.txt

Output Set: N:\CRF3\05112001\I619047C.raw

```

226 Met Gly Lys Ser Gly Lys Ser Gly Asn Gly Ser Tyr Arg Leu Leu Asp
227      85      90      95
229 cat tac aaa tat ctt act gca tgg ttt gaa ctt ctt aat tta cca aag      336
230 His Tyr Lys Tyr Leu Thr Ala Trp Phe Glu Leu Leu Asn Leu Pro Lys
231      100      105      110
233 aag atc att ttt gtc ggc cat gat tgg ggt gct tgt ttg gca ttt cat      384
234 Lys Ile Ile Phe Val Gly His Asp Trp Gly Ala Cys Leu Ala Phe His
235      115      120      125
237 tat agc tat gag cat caa gat aag atc aaa gca ata gtt cac gct gaa      432
238 Tyr Ser Tyr Glu His Gln Asp Lys Ile Lys Ala Ile Val His Ala Glu
239      130      135      140
241 agt gta gta gat gtg att gaa tca tgg gat gaa tgg cct gat att gaa      480
242 Ser Val Val Asp Val Ile Glu Ser Trp Asp Glu Trp Pro Asp Ile Glu
243 145      150      155      160
245 gaa gat att gcg ttg atc aaa tct gaa gaa gga gaa aaa atg gtt ttg      528
246 Glu Asp Ile Ala Leu Ile Lys Ser Glu Glu Gly Glu Lys Met Val Leu
247      165      170      175
249 gag aat aac ttc ttc gtg gaa acc atg ttg cca tca aaa atc atg aga      576
250 Glu Asn Asn Phe Phe Val Glu Thr Met Leu Pro Ser Lys Ile Met Arg
251      180      185      190
253 aag tta gaa cca gac gaa gtt gac gca tat ctt gaa cca ttc aaa gag      624
254 Lys Leu Glu Pro Asp Glu Val Asp Ala Tyr Leu Glu Pro Phe Lys Glu
255      195      200      205
257 aaa ggt gaa gtt cgt cgt cca aca tta tca tgg cct cgt gaa atc ccg      672
258 Lys Gly Glu Val Arg Arg Pro Thr Leu Ser Trp Pro Arg Glu Ile Pro
259      210      215      220
261 tta gta aaa ggt ggt aaa cct gac gtt gta caa att gtt agg aat tat      720
262 Leu Val Lys Gly Gly Lys Pro Asp Val Val Gln Ile Val Arg Asn Tyr
263 225      230      235      240
265 aat gct tat cta cgt gca agt gat gat tta cca aaa atg ttt att gaa      768
266 Asn Ala Tyr Leu Arg Ala Ser Asp Asp Leu Pro Lys Met Phe Ile Glu
267      245      250      255
269 tcg gat cca gga ttc ttt tcc aat gct att gtt gaa ggc gcc aag aag      816
270 Ser Asp Pro Gly Phe Phe Ser Asn Ala Ile Val Glu Gly Ala Lys Lys
271      260      265      270
273 ttt cct aat act gaa ttt gtc aaa gta aaa ggt ctt cat ttt tcg caa      864
274 Phe Pro Asn Thr Glu Phe Val Lys Val Lys Gly Leu His Phe Ser Gln
275      275      280      285
277 gaa gat gca cct gat gaa atg gga aaa tat atc aaa tcg ttc gtt gag      912
278 Glu Asp Ala Pro Asp Glu Met Gly Lys Tyr Ile Lys Ser Phe Val Glu
279      290      295      300
281 cga gtt ctc aaa aat gaa caa taa      936
282 Arg Val Leu Lys Asn Glu Gln
283 305      310
286 <210> SEQ ID NO: 4
287 <211> LENGTH: 311
288 <212> TYPE: PRT
289 <213> ORGANISM: Renilla reniformis (mutated sequence)
291 <400> SEQUENCE: 4

```

## RAW SEQUENCE LISTING

DATE: 05/11/2001

PATENT APPLICATION: US/09/619,047C

TIME: 13:55:55

Input Set : A:\CHEM1110.ST25.txt

Output Set: N:\CRF3\05112001\I619047C.raw

```

293 Met Thr Ser Lys Val Tyr Asp Pro Glu Gln Arg Lys Arg Met Ile Thr
294 1 5 10 15
297 Gly Pro Gln Trp Trp Ala Arg Cys Lys Gln Met Asn Val Leu Asp Ser
298 20 25 30
301 Phe Ile Asn Tyr Tyr Asp Ser Glu Lys His Ala Glu Asn Ala Val Ile
302 35 40 45
305 Phe Leu His Gly Asn Ala Ala Ser Ser Tyr Leu Trp Arg His Val Val
306 50 55 60
309 Pro His Ile Glu Pro Val Ala Arg Cys Ile Ile Pro Asp Leu Ile Gly
310 65 70 75 80
313 Met Gly Lys Ser Gly Lys Ser Gly Asn Gly Ser Tyr Arg Leu Leu Asp
314 85 90 95
317 His Tyr Lys Tyr Leu Thr Ala Trp Phe Glu Leu Leu Asn Leu Pro Lys
318 100 105 110
321 Lys Ile Ile Phe Val Gly His Asp Trp Gly Ala Cys Leu Ala Phe His
322 115 120 125
325 Tyr Ser Tyr Glu His Gln Asp Lys Ile Lys Ala Ile Val His Ala Glu
326 130 135 140
329 Ser Val Val Asp Val Ile Glu Ser Trp Asp Glu Trp Pro Asp Ile Glu
330 145 150 155 160
333 Glu Asp Ile Ala Leu Ile Lys Ser Glu Glu Gly Glu Lys Met Val Leu
334 165 170 175
337 Glu Asn Asn Phe Phe Val Glu Thr Met Leu Pro Ser Lys Ile Met Arg
338 180 185 190
341 Lys Leu Glu Pro Asp Glu Val Asp Ala Tyr Leu Glu Pro Phe Lys Glu
342 195 200 205
345 Lys Gly Glu Val Arg Arg Pro Thr Leu Ser Trp Pro Arg Glu Ile Pro
346 210 215 220
349 Leu Val Lys Gly Gly Lys Pro Asp Val Val Gln Ile Val Arg Asn Tyr
350 225 230 235 240
353 Asn Ala Tyr Leu Arg Ala Ser Asp Asp Leu Pro Lys Met Phe Ile Glu
354 245 250 255
357 Ser Asp Pro Gly Phe Phe Ser Asn Ala Ile Val Glu Gly Ala Lys Lys
358 260 265 270
361 Phe Pro Asn Thr Glu Phe Val Lys Val Lys Gly Leu His Phe Ser Gln
362 275 280 285
365 Glu Asp Ala Pro Asp Glu Met Gly Lys Tyr Ile Lys Ser Phe Val Glu
366 290 295 300
369 Arg Val Leu Lys Asn Glu Gln
370 305 310

```

373 &lt;210&gt; SEQ ID NO: 5

374 &lt;211&gt; LENGTH: 8

375 &lt;212&gt; TYPE: PRT

376 &lt;213&gt; ORGANISM: Artificial sequence

378 &lt;220&gt; FEATURE:

379 &lt;223&gt; OTHER INFORMATION: Description of Artificial sequence: Protease recognition sequence

381 &lt;400&gt; SEQUENCE: 5

383 Ser Gln Asn Tyr Pro Ile Val Gln

384 1 5

**Please Note:**

Use f n and/ r Xaa have been detected in the Sequenc Listing. Please review the Sequence Listing t nsure that a c rresp nding explanati n is presented in the <220> to <223> fields of each sequence which presents at least ne n or Xaa.

VERIFICATION SUMMARY

DATE: 05/11/2001

PATENT APPLICATION: US/09/619,047C

TIME: 13:55:56

Input Set : A:\CHEM1110.ST25.txt

Output Set: N:\CRF3\05112001\I619047C.raw

L:613 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22  
L:632 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23  
L:677 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26